Journal of Regulatory Economics; Mar 2003; 23, 2; ABI/INFORM Global
pg. 193

Journal of Regulatory Economics; 23:2 193–205, 2003

Using Random *n*th Price Auctions to Value Non-Market Goods and Services*

JOHN A. LIST
University of Maryland
AREC, 2200 Symons Hall, College Park, MD 20742-5535
E-mail: jlist@arec.umd.edu

Abstract
Public policy decision making often requires balancing the benefits of a policy with the costs. While regulators in the United States and abroad rely heavily on benefit-cost analysis, critics contend that hypothetical bias precludes one of the most popular benefit estimation techniques—contingent surveys—from providing reliable economic values for nonmarket goods and services. This paper explores a new methodology to obtain the total value of nonmarket goods and services—random *n*th price auctions. The empirical work revolves around examining behavior of 360 participants in a competitive marketplace, where subjects naturally buy, sell, and trade commodities. The field experiment provides some preliminary evidence that hypothetical random *n*th price auctions can, in certain situations, reveal demand truthfully.

1. Introduction

Benefit-cost analysis remains the central paradigm used throughout the public sector. While regulators and policymakers around the world use benefit-cost analysis, properly estimating the total benefits of nonmarket goods and services has proven quite difficult. Although the contingent valuation method (CVM) remains the most popular method to estimate the total economic value of the commodity in question, unfortunately it has not been without its shortcomings: in practice, many studies find that hypothetical statements

* I would like to thank Michael Crew and two anonymous reviewers for very helpful comments. Without indicting, I should note that Catherine Eckel's contribution to the development of this manuscript was invaluable. Richard Carson, Vernon Smith, and Laura Taylor provided useful suggestions during the research process. The usual caveats apply.

of value exceed actual values in CVM markets.[1] An overstatement of value is of considerable importance since credibly valuing nonmarket goods and services remains critical for numerous reasons: if contingent markets are used to estimate values, and hypothetical bias permeates stated economic values, then, for example, too many proposals will pass a benefit-cost test, the world's resources will be overvalued—making green accounting procedures illegitimate—and plaintiffs will receive too much in settlement damages.

With hypothetical bias in mind, the National Oceanic and Atmospheric Administration's (NOAA) blue-ribbon panel, composed of icons in the profession such as Kenneth Arrow and Robert Solow, recommended that hypothetical values be adjusted by a "divide by 2" rule unless the values can be calibrated using actual market data (NOAA 1994, 1996). Given that the "divide by two" rule has been considered quite ad hoc, the NOAA rule has served to motivate more research into the nature of calibrating hypothetical and actual values.

In this study, I explore a relatively new methodology to obtain total economic values for nonmarket goods and services—random *n*th price auctions. Akin to Vickrey's second-price auction, the random *n*th price auction is theoretically incentive compatible. While the random *n*th price auction has firm roots theoretically, to gain widespread acceptability it must be scrutinized under the same rigorous tests that other valuation institutions have been subjected to (e.g., Cummings et al. 1995, 1997). While these other valuation institutions, such as open-ended and dichotomous choice questions, Vickrey second-price auctions, and the Becker-DeGroot-Marschak mechanism have generally not performed well, this study represents a first attempt at providing a firm understanding of the external validity properties of random *n*th price auctions.

In the empirical analysis, I compare hypothetical and actual bids from random *n*th price auctions (as well as from Vickrey second-price auctions) to explore the auction's properties in an actual marketplace. Since one nice characteristic of the random *n*th price auction is its natural ability to engage bidders even if they believe they are not near the upper tail of the value distribution, one might expect that its performance will differ from that of a standard Vickrey second-price auction. Overall, I find results consonant with this notion. Examining bidding behavior of 360 subjects in a competitive marketplace, I find that hypothetical and actual bid distributions from random *n*th price auctions are more similar than comparable bid distributions from Vickrey's second-price auction. Most notably, in the actual auctions the lower tail of bids obtained from the Vickrey auction is significantly below the lower tail of bids obtained from the random *n*th price auction, inducing a larger wedge between hypothetical and actual values in the Vickrey auction. This result represents an important advance in the nonmarket valuation literature, where to date few mechanisms have elicited equivalent actual and hypothetical values.

The remainder of this study proceeds as follows. Section 2 presents some background

---

1 I should note that studies have not always found hypothetical values that exceed actual values (e.g., List 2001b). The interested reader should read the debate between Cummings et al. (1997) and Haab et al. (1999) and Smith (1999). Table 1 below provides a summary of willingness to pay studies that examine the disparity.

Table 1. Summary of Select WTP Studies

Study	Year	Type of Experiment	Type of Good	Type of Elicitation	Calibration Factor[a]
Bohm	1972	Laboratory	Private	Open-ended	1.00
					1.16
					1.16
					1.34
Bishop and Heberlein	1979	Field	Private	Dichotomous	0.30–1.60
Bishop and Heberlein	1986	Field	Private	Open-ended	1.30–2.30
				Dichotomous	0.80
Brookshire and Coursey	1987	Field and lab	Public	Smith	2.00
					1.85
Dickie et al.	1987	Field	Private	Dichotomous	1.00
Coursey et al.	1987	Laboratory	Private	Vickrey	1.00
Sinden et al.	1988	Laboratory	Public	Open-ended	0.80–1.50
Kealy et al.	1988	Laboratory	Private	Dichotomous	1.40
			Public		1.40
Kealy et al.	1990	Laboratory	Private	Dichotomous	1.30
			Public		1.00–2.00
Seip and Strand	1992	Field	Public	Dichotomous	10.30
Navrud	1992	Field	Private	Dichotomous	3.20
			Public		1.60–2.10
Irwin et al.	1992	Laboratory	Private	Vickrey	1.00
					2.50
Boyce et al.	1992	Field and lab	Private	BDM	1.50
					2.10
					0.90
McClelland et al.	1993	Laboratory	Private	Vickrey	2.20
					0.80
Neill et al.	1994	Laboratory	Private	Open-ended	3.1–25.1
Loomis et al.	1996	Laboratory	Private	Open-ended	1.80–2.90
				Dichotomous	2.00–3.60
Brown et al.	1996	Field	Public	Open-ended	4.10
				Dichotomous	6.50
Frykblom	1997	Laboratory	Private	Dichotomous	1.50
Loomis et al.	1997	Laboratory	Private	Open-ended	1.86
				Dichotomous	2.54–3.00
Cummings et al.	1997	Laboratory	Public	Referenda	[b]
Spencer et al.	1998	Laboratory	Public	Provision pt.	4.67
					4.66
List and Shogren	1998a	Field	Private	Vickrey	2.54
					3.47
					2.19
Fox et al.	1998	Laboratory	Private	Vickrey	1.20
					1.50
Balistreri et al.	1998	Laboratory	Private	Open-ended	1.25
				Dichotomous	1.54
				Dichotomous	0.58
Johannesson et al.	1998	Laboratory	Private	Dichotomous	1.18
					0.80
					1.29
					0.88
Frykblom	1999	Laboratory	Private	Vickrey	1.33
List	2001	Field	Private	Choice Exp.	[b]

[a]Calibration factor is calculated as mean hypothetical/mean actual.
[b]Study gathered "Yes"/ "No" votes rather than actual dollar values.

information and the experimental design. Section 3 provides the empirical results. Section 4 concludes.

2. Background and Experimental Design

The economics discipline as a whole has only within the past few decades begun to recognize the valuable insights that experimental methods can provide. Yet, experimental economics has now become such a fixture within the profession that many feel a Nobel Prize is imminent for the pioneers of the subject. While experimental methods are unique in that they can, for example, cleanly test decision or game theoretic models and examine institutional details and procedures, the thrust of much of my recent experimental work, and the current study, focuses on improving our understanding of important policy problems. In particular, by using experimental methods to examine behavior within well-functioning markets, my work has focused on increasing external validity without unduly sacrificing internal validity.[2]

Gathering primary data from a marketplace to provide insights into policy issues represents a nice complement to laboratory examinations. As I have argued elsewhere (e.g., List 2001a; List and Lucking-Reiley 2000) unlike econometric studies using field data, field experiments represent an exciting opportunity to test the validity and relevance of the predictions of economic theory because one can create exogenous variation in the variables of interest, allowing for relatively novel tests of economic theory. In addition, a field experiment is able to check the robustness of laboratory results in a natural setting, where the mathematical assumptions of the theory cannot necessarily be guaranteed to hold. Hence, controlled field studies provide a useful middle ground between the controlled environment of the laboratory and econometric studies using field data, in which the researcher relies on estimating a correctly specified equation that includes all of the important control variables.[3]

Before becoming immersed in the details of the experimental design, it is important to understand exactly why we should be concerned about contingent markets and how the current experimental treatments can be helpful in the policy arena. A useful starting point is President Reagan's 1981 Executive Order 12291, which requires that federal agencies consider both the benefits and costs of regulations prior to their implementation. While economists have long measured the benefits and costs of private goods routinely bought and sold in the marketplace, a much more difficult task faces the regulator interested in

2 External validity refers to the ability to generalize from the research context to settings that the research is intended to describe. Internal validation refers to the ability to draw conclusions from one's research.

3 I would be remiss not to mention potential drawbacks of field experiments. One particularly salient point is that field experiments may not be as ''clean'' as laboratory experiments, where researchers have greater control: inducing preferences to accord with theoretical assumptions, and excluding other complicating factors. Perhaps a useful compromise between lab experiments run with student subjects and field experiments is a laboratory experiment with ''non-standard'' subject pools. Examples of such subject pools are representatives from the overall population or any given select population (see, for example, List 2001c).

estimating the benefits of increased air and water quality, for example. Since ordinary markets typically do not exist for public goods, such as environmental improvements, sending an astronaut to the moon, etc., federal agencies such as the environmental protection agency (EPA) have an acute need for practical, credible, methods to estimate benefits and costs. While the appropriate recognition of costs is invaluable, estimating the benefits of nonmarket goods and services remains a very controversial issue and is at the center of this study. Although several methods are potentially useful in benefit estimation, it is believed that the flexibility and totality of the CVM approach renders it the ''only game in town'' (Cummings et al. 1986).

The CVM approach is typically based on the simple notion of constructing a market and obtaining participants' maximum willingness to pay for a good or service. Yet, using contingent surveys to estimate values for goods and services generally causes uneasiness. This discomfort has translated into able scholars questioning the entire CVM paradigm, typically asking whether a number obtained from a contingent survey is better than no number at all (Diamond and Hausman 1994). Unquestionably, part of the uneasiness is due to the fact that the CVM approach is hypothetical in nature: hypothetical markets, hypothetical commodity provisioning, and hypothetical payment, for example. Accordingly, the persistence that ''hypothetical bias'' precludes contingent markets from providing reliable value estimates is understandable.

Prior to moving on, it is important to understand exactly what ''hypothetical bias'' means in the literature and why it is important for regulators to understand. One definition would be as follows: hypothetical bias is the difference between hypothetical and actual statements of value, where actual statements of value are obtained from experiments with real economic commitments. As such, implicitly the literature assumes that cash-based estimates are unbiased signals of preferences. A first important question, therefore, is whether the various valuation institutions used in the lab (e.g., open-ended and dichotomous choice questions, Vickrey second-price auctions, and the Becker-DeGroot-Marschak mechanism) are incentive compatible? Or, likewise, do these mechanisms elicit true values? This question can naturally be answered via experimental methods, as experimentalists have the necessary tools to examine incentive compatibility.[4] Running the risk of over-generalizing, there is some evidence that suggests open-ended valuation questions and Vickrey's second-price auction do not universally perform well (see discussion below).

A second important question concerns actual versus hypothetical statements of value from the same institution. This relationship is important because for many nonmarket goods and services the researcher will only be in a position to hypothetically provide the good, and therefore an understanding of the hypothetical/actual relationship with goods that can actually be provided represents a useful first step to examine contingent valuation. I provide table 1, which extends List and Shogren (2002), as a brief summary of the recent literature that compares hypothetical and actual willingness to pay responses. In columns 2–4 of table 1, I have attempted to highlight differences across the various studies that the

4 One simple method to accomplish this task is to use ''induced values'' (each subject receives a private resale value for the good) and compare elicited values (e.g., bids) to resale values.

researcher can control. For example, important methodological features across these studies include the setting of the experiment (laboratory, field, or both), type of good (public or private), and elicitation method (open-ended, Vickrey 2nd price auction, dichotomous choice, provision point mechanism, Smith auction, referendum, choice experiments, or Becker-DeGroot-Marschak (BDM)).

As one can glean from the fourth column of table 1, most of the studies in the literature use theoretically incentive compatible mechanisms (all but open-ended surveys are theoretically incentive compatible). And, within the rather large literature that uses theoretically incentive compatible mechanisms, results in the rightmost column suggest that the average subject tends to provide hypothetical statements of value that exceed actual values. This nuance is of considerable importance: if contingent markets are used to estimate economic values, then hypothetical bias will preclude regulators from carrying out unbiased benefit-cost analyses and green accounting. This evidence of overstatement, coupled with the aforementioned NOAA decree, has triggered a search for an elicitation technique that yields comparable value estimates across hypothetical and actual statements of value.

This study examines an alternative institution to compare hypothetical and actual willingness to pay statements: random *n*th price auctions. While the random *n*th price auction shares certain theoretic properties with Vickrey's second-price mechanism, some studies have observed that in practice second-price auctions may not be successful at engaging all bidders to truthfully reveal values. For example, Shogren et al. (2001) found that bidders with values near the likely second-price bid sincerely, but those with values well below the price often submit bids below their true values.[5] This behavior persists when insincere bids remain undetected and unpunished by the market. Accordingly, the difference in hypothetical and actual values found in previous Vickrey second-price auction studies (and perhaps studies using other valuation institutions) may be systematically due to bias in *both* the hypothetical and actual elicitations, rather than purely hypothetical bias. If random *n*th price auctions can attenuate one or both of these biases, then regulators are one step closer to more efficient policy design.

2.1. Experimental Design

The field experiment, which was conducted at a sportscard show in Tucson, AZ, is designed to provide an initial examination of hypothetical and actual bidding behavior within random *n*th price auctions.[6] Table 2 presents a summary of the 2X6 experimental design. Each of the six treatments (Vickrey second-price and random *n*th price auctions: hypothetical, hypothetical with cheap talk, or actual with consumers and dealers) is essential for a thorough study of hypothetical bias. A comparison between bids obtained in

5 This result suggests that the Vickrey auction may be useful in gathering marginal (equilibrium) values, but since the elicited demand curve may not be accurate, gathering total values via the Vickrey auction may be quite misleading.

6 Note that List and Shogren (2002) examine willingness to accept (seller's) behavior within a random *n*th price auction. They find evidence that people understate their unconditional real willingness to accept in hypothetical scenarios. But, after controlling for person-specific effects, the observed difference vanishes.

Table 2. Experimental Design

Subject Type	Random *n*th and Second-Price Auctions	Random *n*th and Second-Price Auctions	Random *n*th and Second-Price Auctions
Nondealers	Hypothetical	Hypothetical with cheap talk	Actual
Dealers	Hypothetical	Hypothetical with cheap talk	Actual

Notes. Each cell represents two unique treatments. For example, "Hypothetical" in row 1, column 1 denotes that one treatment had nondealers competing in a random *n*th price hypothetical auction and another had a *different* group of nondealers competing in a hypothetical second-price auction. No subject participated in more than one treatment. In each auction, I auctioned-off a 1983 Cal Ripken Jr. *Topps* professionally graded baseball card.

treatments displayed in column 1 versus column 3 provides the standard hypothetical versus actual valuation comparison that studies reported in table 1 have examined. To supplement these treatments, I include an extra set of treatments that measure whether hypothetical statements of value are influenced by using a "cheap talk" script (see, for example, Cummings and Taylor 1999). The cheap talk script, which is described more fully below, is an attempt to engage subjects in the cognitive process of valuing the good even though they do not have a monetary incentive to be engaged. The script has had success amongst certain subject pools in deflating hypothetical bids to more closely approximate actual values (Cummings and Taylor 1999; List 2001a).

In practice, each of the six treatments had four steps. Step 1—as a potential subject entered the floor of the sportscard show, I asked if she was interested in the sportscard on the table. If she answered in the affirmative, I proceeded to provide a thorough description of the sportscard's condition, ensuring that there was no ambiguity about the card's grade. For all six auction types, I chose a Cal Ripken, Jr. 1983 *Topps* baseball card, which has a book value of approximately $12. All auctions displayed the same sportscard to all bidders—a Cal Ripken Jr. PSA graded "PSA 8 near mint/mint" baseball card. The monitor worked one-on-one with the participant and no time limit was imposed on her inspection of the card.

In Step 2, the administrator gave the participant an instruction sheet that consisted of two parts: (1) an auction rules sheet that also included an illustrative example of the appropriate auction, and (2) a bidding sheet.[7] In the Vickrey actual auctions, subjects were informed that the winner of the card would pay a price equal to the amount of the second-highest bid. In the actual random *n*th price treatments, subjects were informed that winner(s) of the card would pay a price equal to the amount of the *n*th highest bid. The auction rules for the random *n*th-price auction were stated in 4 simple steps: (i) each bidder submits a bid; (ii) each bid is rank-ordered; (iii) the monitor selects a random number (n) uniformly-distributed between 2 and Z (Z bidders); and (iv) the monitor sells one unit of the good to each of the (n-1) highest bidders at the *n*th-price (where a different *n* is drawn for each treatment). For the hypothetical auctions, I follow the nomenclature of previous

7 All experimental instructions are available upon request.

studies and state (in the random *n*th price case): ''suppose you were to bid on the sportscard on the table, if the winner(s) of the card were to pay a price equal to the amount of the *n*th highest bid, how much would you bid?'' Subjects participating in the hypothetical with cheap talk auctions read the cheap talk script just prior to placing their hypothetical bid. The cheap talk script is similar to List (2001a), and included the following:

> In some recent studies, several different groups of people participated in a hypothetical survey. No one had to pay money, and the results of these studies were that on average, across the groups, people overstated their actual willingness to pay.
> We call this ''hypothetical bias.'' Hypothetical bias is the difference that we continually see in the way people respond to hypothetical situations as compared to real situations . . .
> How can we get people to think about their true values in a hypothetical exercise like they think in a real exercise? How do we get them to think about what it means to really dig into their pocket and pay money, if in fact they really aren't going to have to do it? Let me tell you why I think that we continually see this hypothetical bias, why people behave differently in a hypothetical scenario. I think that when we hypothetically value something we ''guess'' and do not put forth effort rather than really try to provide a true answer. This is just my opinion, of course, but it's what I think may be going on in hypothetical surveys. So if I was in your shoes I would choose as if you were really going to face the consequences of your choice. Please keep this in mind in your decisions.[8]

After all questions were answered, each participant placed a bid on the bidding sheet (Step 3). In Step 4, the administrator explained that if the participant won the auction, she would be contacted by email or telephone within three days (for those in the actual auctions). Upon receipt of payment, I would send her the card (postage paid).[9]

Given that the cheap talk design has not been found to work across all subject pools (List 2001a), an added advantage of experimenting with bidder type (e.g., conducting some of the treatments with professional card dealers and others with nondealers) is that it provides evidence on whether the cheap talk script universally influences values. In each

8 Note that it is quite possible that my cheap talk script induces certain biases. For example, it could merely shift down an already accurate demand curve rather than make it more accurate; or it could induce an overreaction. A comparison of the bidding data in the actual treatment with data from the cheap talk treatment allows a test of this hypothesis (see below).

9 Two aspects of my design make the current experimental construct somewhat different from ''typical'' lab experiments: (i) subjects are not paid for their participation, and (ii) subjects do not know the experimental results for three days. There is some evidence that the ''found money'' effect is not important in experiments (see, e.g., List 2001c). Yet, whether subjects' field behavior is affected by the time lag associated with the announcement of winners is unknown. While I cannot address this issue directly within the context herein, there is some evidence that suggests delay of payment does influence behavior in the lab (Coller and Williams 1999). Whether this behavior spills-over to the field is open for debate, but should be the subject of future research.

Table 3. Summary Statistics for Vickrey and Random *n*th Price Auctions[a,b]

	Dealers			Non-Dealers		
	Hyp. Auction $N = 30$	Hyp. with Cheap Talk $N = 30$	Actual Auction $N = 30$	Hyp. Auction $N = 30$	Hyp. with Cheap Talk $N = 30$	Actual Auction $N = 30$
Vickrey Mean Bid	\$6.67 (\$5.08)	\$5.05 (\$3.11)	\$2.28 (\$2.07)	\$7.40 (\$6.53)	\$2.58 (\$2.99)	\$2.78 (\$3.11)
t-test of means hyp. versus	—	$t = -1.50$	$t = -4.38^*$	—	$t = -3.67^*$	$t = -3.50^*$
t-test of means hyp. with cheap talk versus	$t = -1.50$	—	$t = -4.07^*$	$t = -3.67^*$	—	$t = -0.25$
Random *n*th Mean Bid	\$7.18 (\$6.24)	\$4.97 (\$3.16)	\$3.67 (\$2.38)	\$8.65 (\$9.57)	\$3.54 (\$3.35)	\$3.42 (\$2.21)
t-test of means hyp. versus	—	$t = -1.73$	$t = -2.88^*$	—	$t = -2.76^*$	$t = -2.92^*$
t-test of means hyp. with cheap talk versus	$t = -1.73$	—	$t = -1.80$	$t = -2.76^*$	—	$t = 0.16$

[a] Standard deviations are in parentheses beneath means.
[b] *t*-tests are large sample *t*-tests.
* Significantly different values at the two-tailed $p < 0.05$ level.

case, I randomize participants into *one* of the six bidding formats to assure that subpopulations are similar. For example, while gathering the nondealer data, at the top of each hour I changed treatments, and continued to change treatments until each cell had 30 observations. For dealers, I gathered data before the show began and alternated treatment type until each treatment had 30 observations. Again, no subjects participated in more than one treatment.

3. Experimental Results

The general empirical analysis will proceed by examining the field data and when behavior differs, I analyze whether institutional considerations, general market experience, or other demographic factors might be responsible for the observed behavioral differences. Empirical results are contained in table 3. Table 3 summarizes the Vickrey second-price auction results in the upper panel and the random *n*th price auction results in the lower panel. Each auction type includes responses from 90 dealers and 90 nondealers, equally distributed across the three treatments. In general, the Vickrey auction data are consistent with List (2001a) and suggest a few key insights. First, in the dealer treatment data displayed in columns 1–3, I find that average bids in both hypothetical regimes are greater than the average actual bid ($6.67 and $5.05 versus $2.28). In the non-dealer treatments reported in columns 4–6, mean bids in the actual and hypothetical with cheap talk auctions are similar, but much lower than the mean hypothetical bid ($2.78 and $2.58 versus $7.40). The ratio of mean hypothetical to mean real is nearly 3 across both dealers and nondealers. This degree of overstatement is in the range of values reported in table 1.

The second and third rows in the top panel of table 3 contain parametric t-tests for the equivalency of mean bids across auction types: $H_o: \bar{x}_i = \bar{x}_j$, where $\bar{x}$ represents mean bid. Since there are three auction types in each group, I compute three distinct t-tests. Table 3 can be read as follows: the dealer two-tailed t-test of hypothetical versus actual is at the intersection of row 2, column 3, and indicates $t = -4.38$, which suggests the two means are significantly different at the $p < 0.01$ level. In the dealer auctions, note from row 2, column 2, that hypothetical bids are not statistically different from hypothetical with cheap talk bids at conventional significance levels ($t = -1.50$). Conversely, t-tests in rows 2 and 3 indicate that the mean bid from the actual auction <u>is</u> statistically different from mean bids in the hypothetical $t = -4.38$ and hypothetical with cheap talk ($t = -4.07$) auctions at the $p < 0.01$ level. To provide a robustness test, I also examine non-parametric statistical tests. Non-parametric Mann-Whitney tests of whether the sampled populations have identical probability distributions reveal the same insights: the distribution of actual bids is to the left of both hypothetical bid distributions, while the hypothetical distributions are located similarly. These results suggest that the cheap talk design failed to eliminate hypothetical bias in the dealer valuation exercises. This finding is consonant with List (2001a), who auctioned off a $250 sportscard.

Data from the non-dealer Vickrey treatments contrast with data in the dealer auctions: t-tests suggest the mean hypothetical bid is statistically different from mean bids in the hypothetical with cheap talk and actual auctions at the $p < 0.01$ level ($t = -3.67$; $t = -3.50$). More importantly, the average actual bid is not statistically different from the average hypothetical with cheap talk bid at conventional levels ($t = -0.25$). Mann-Whitney tests reveal similar insights: actual and hypothetical with cheap talk bids are significantly less than bids in the hypothetical regime. Consistent with List (2001a), these results suggest that the hypothetical with cheap talk design effectively eliminated hypothetical bias for ordinary consumers.

The lower panel in table 2 contains a summary of the nondealer and dealer data obtained from the random nth price auctions, and again includes responses from 180 subjects, equally distributed across the various treatment types. These results represent a first attempt at comparing hypothetical and actual bids elicited via a random nth price auction. Akin to results in the Vickrey auctions, dealers tend to bid higher in the hypothetical auctions compared to the actual auction ($7.18 and $4.97 versus $3.67). But, in this case, the comparison of bids in the actual and hypothetical with cheap talk auctions suggests that the two distributions are not different from one another at conventional levels, although it is important to recognize that both parametric and nonparametric tests indicate the distributions are significantly different at the $p < 0.10$ level (e.g., $t = -1.80$). Given this nearly significant result, one should be hesitant to conclude firmly that the bid distributions are identical, as Type II error should be considered, but this result does suggest that the hypothetical and actual bid distributions from random *n*th price auctions are more similar than comparable bid distributions from Vickrey's second price auction.

This convergence in bid distributions occurs because of an increase in the average actual bidding level in the random *n*th price auction (Vickrey mean = \$2.28; Random *n*th mean = \$3.67). This difference in means is largely due to an increase in the lower tail of bids in the random *n*th price auction: whereas 25/30 (13/30) dealers bid below \$5 (\$2) in the second-price auction, only 18/30 (4/30) dealers bid below \$5 (\$2) in the random *n*th price auction. To test if these differences are statistically significant, I use a Fisher's exact test, which has a hypergeometric distribution under the null. Results of the exact tests suggest that the homogeneity null hypothesis should be rejected at the $p < 0.05$ level for both subsamples (\$5: $z = 2.00$; \$2: $z = 2.58$), implying that significantly more bidders place bids below \$5 and \$2 in the Vickrey second-price auction compared to the random *n*th price auction. These results are in line with those from the induced value experimental results of Shogren et al. (2001), yet herein I cannot formally examine whether bidders are submitting bids below or above their true valuations since I am using homegrown, rather than induced, values.[10]

Overall, data from the nondealer random *n*th price auction treatments are consistent with experimental results in the Vickrey second-price auctions—the hypothetical with cheap talk design effectively eliminated hypothetical bias. Yet, similar to bidding behavior in the actual dealer data, there is evidence that low valued consumers place larger bids in random *n*th price auction, leading to the finding that the average bid is larger in the auctions where every bidder can potentially be included. For instance, whereas 21/30 (13/30) nondealers bid below \$5 (\$2) in the Vickrey second-price auction, only 16/30 (8/30) nondealers bid below \$5 (\$2) in the random *n*th price auction. While these tendencies are consistent with the dealer data, a Fisher's Exact test suggests that the null hypothesis cannot be rejected at conventional levels (\$5: $z = 1.32$; \$2: $z = 1.35$).[11]

---

10 A theory that can potentially shed light on this finding is due to Smith and Walker (1993) and Smith and Szidarovszky (1999), who present effort models which predict that individuals' behavior will more closely match the predictions of rational-behavior theories as the expected stakes of the decision increase.

11 Although analysis of the raw data provides interesting insights, there has been no attempt to control for other factors that may influence the individual bidding level. As in any well-designed experiment that is interested in behavioral differences across institutions, a regression analysis is unnecessary since subjects are randomly allocated into treatments. As aforementioned, in the current field experiment, subjects were randomly inserted into *one* of the six bidding formats to assure that sub-populations are similar. Yet, if for some reason subjects arrived nonrandomly at my dealer table (e.g., high income subjects had to take their daughters to a t-ball game and could not arrive until 3 pm), it is important to examine whether the results are robust to controlling for subject-specific factors, such as years of trading experience, gender, income, education, and age. To condition on these factors, I estimate the following bid regression model: $bid = g(\alpha + \beta' X)$; where X includes years of trading experience, yearly income, age, gender, education, and dichotomous variables indicating treatment type and bidder type. Empirical results from this exercise are consonant with inference gained from the raw data.

4. Concluding Remarks

Whether hypothetical statements map into real behavior merits serious consideration since contingent markets represent the only viable technique to estimate the total value of nonmarket goods and services. Since credible value estimation permeates several important policy arenas—from damage assessment to everyday benefit-cost analysis—developing a reliable method to obtain economic values remains critical. In this study, I examine individual bidding behavior across 360 subjects in a well-functioning marketplace as a first attempt at providing an understanding of the external validity properties of random *n*th price auctions. The initial results are promising. In most cases, hypothetical values gathered via a cheap talk treatment are statistically indistinguishable from actual responses.

The evidence herein represents a useful first step toward credible value estimation. Yet, there remains much work to be done. For example, while the results suggest that the random *n*th price auction may work for valuing private goods, do these results carry over to public goods? Also, are these findings directly related to the subject pool, context, or nature of the private good? The answers to these questions are empirical in nature and not immediately obvious. While these particular questions can eventually be answered via experimental methods, an equally important goal should be a conversion of the body of experimental results into theory. The debate on the validity of CVM will likely be palliated only when a robust theoretical or behavioral reason emerges as to why the hypothetical/actual disparity occurs and whether it can be controlled systematically.

References

Coller, M., and M. Williams. 1999. "Eliciting Individual Discount Rates." *Experimental Economics* 2(2): 107–127.

Cummings, R., D. Brookshire, and W. Schulze. 1986. *Valuing Environmental Goods: An Assessment of the Contingent Valuation Method.* Totowa, NJ: Rowman and Allanheld.

Cummings, R. D., G. Harrison, and E. Elisabet Rutström. 1995. "Homegrown Values and Hypothetical Surveys: Is the Dichotomous Choice Approach Incentive Compatible?" *American Economic Review* 85(1): 260–266.

Cummings, R. D., S. Elliott, G. Harrison, and J, Murphy. 1997. "Are Hypothetical Referenda Incentive Compatible?" *Journal of Political Economy* 105: 609–621.

Cummings, R. D., and L. Taylor. 1999. "Unbiased Value Estimates for Environmental Goods: A Cheap Talk Design for the Contingent Valuation Method." *American Economic Review* 89(3): 649–665.

Diamond, P., and J. Hausman. 1994. "Contingent Valuation: Is Some Number Better than no Number?" *Journal of Economic Perspectives* 8: 45–64.

Haab, T. C., J.-C. Huang, and J. C. Whitehead. 1999. "Are Hypothetical Referenda Incentive Compatible? A Comment." *Journal of Political Economy* 107: 186–196.

List, J. A. 2001a. "Do Explicit Warnings Eliminate the Hypothetical Bias in Elicitation Procedures? Evidence from Field Auctions for Sportscards." *American Economic Review* 91(5): 1498–1507.

List, J. A. 2001b. "Using Choice Experiments to Value Non-Market Goods and Services." Working paper, University of Maryland.

List, J. A. 2001c. "Using Public Goods and Prisoner's Dilemma Games to Explore Social Preferences in the Field." *Economic Journal*, Forthcoming.

List, J. A., and D. Lucking-Reiley. 2000. "Demand Reduction in Multi-Unit Auctions: Evidence from a Sportscard Field Experiment." *American Economic Review* 90(4): 961–972.

List, J. A., and J. Shogren. 2002. ''Calibration of Willingness to Accept.'' *Journal of Environmental Economics and Management* 43(2): 219–233.

National Oceanic and Atmospheric Administration, 1994, Natural Resource Damage Assessment: Proposed Rules, Federal Register, 4 May, 59, 23098–23111.

National Oceanic and Atmospheric Administration, 1996, Natural Resource Damage Assessments: Final Rules, Federal Register, 5 January, 61: 439.

Shogren, J., J. A. List, and M. Margolis. 2001. ''A Random *n*th-price Auction.'' *Journal of Economic Behavior and Organization* 46 (4): 409–421.

Smith, V. K. 1999. ''Of Birds and Books: More on Hypothetical Referenda—Comment.'' *Journal of Political Economy* 107: 197–200.

Smith, V. L., and J. M. Walker. 1993. ''Monetary Rewards and Decision Cost in Experimental Economics.'' *Economic Inquiry* 31: 245–261.

Smith, V. L., and F. Szidarovszky. 1999. ''Monetary Rewards and Decision Cost in Strategic Interactions.'' Working paper. University of Arizona.